ML in the Routers: Learn from and Act on Network Traffic

Bing Liu

@NMLRG, ietf95, April 2016

Contents

- A Router’s role in Network Traffic ML
  - Router-involved Learning
  - **In-router Learning** *(our focus)*
- Perception and Future Plan

Router-involved Learning

- Routers participant in the learning process

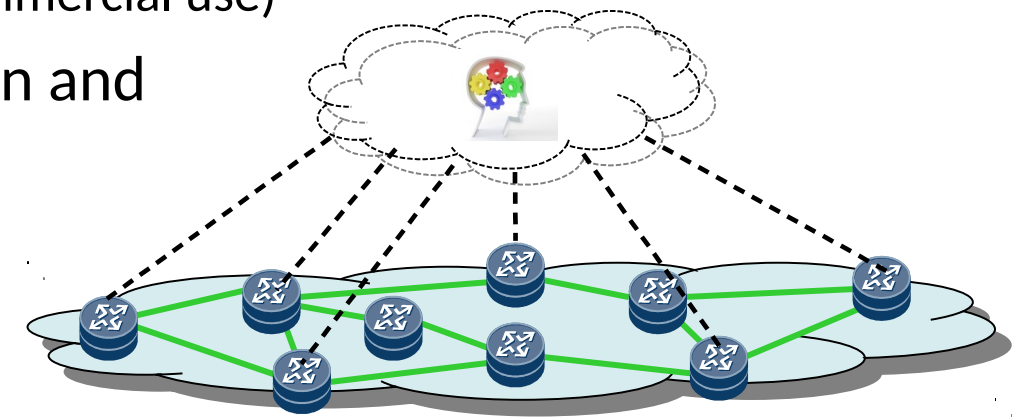
  #1 They act as training data source

  #2 They do the decision/prediction according to the learnt pattern/rules

- But they don’t do the training/learning
- Learning is mostly done in a central entity

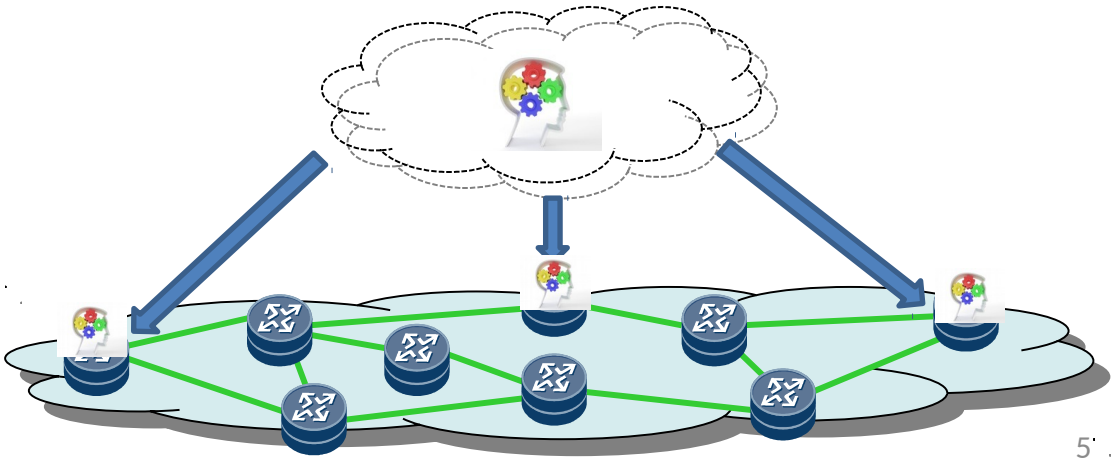
#1: Router as a Reporter/Sensor

- Only acts as data source for the (central) ML applicati on, reporting:
  - Traffic statistics: interface counter, IPFIX/NetFlow etc.
  - Router status: hardware status, device profiles etc.
  - etc.
- Use Cases:
  - Anti-DDoS (already in commercial use)
  - Network QoS estimation and optimization
  - etc.



#2: Router as a Decision/Prediction Agent

- The training is done in the Data Center/Cloud/SDN Controller/ NMS server etc. *(Off-line process)*
- The routers apply the learnt rules to the traffic pass through t hem. *(On-line process)*
- Implication
  - One single router has the same Input/Output form as the training data in DC/Cloud/SDN Controller/NMS server etc.
- Use Case:
  - DPI
  - etc.



In-Router Learning

- Routers can learn from the traffic/local data to gain so me knowledge or decision making capability *(predicti on)*

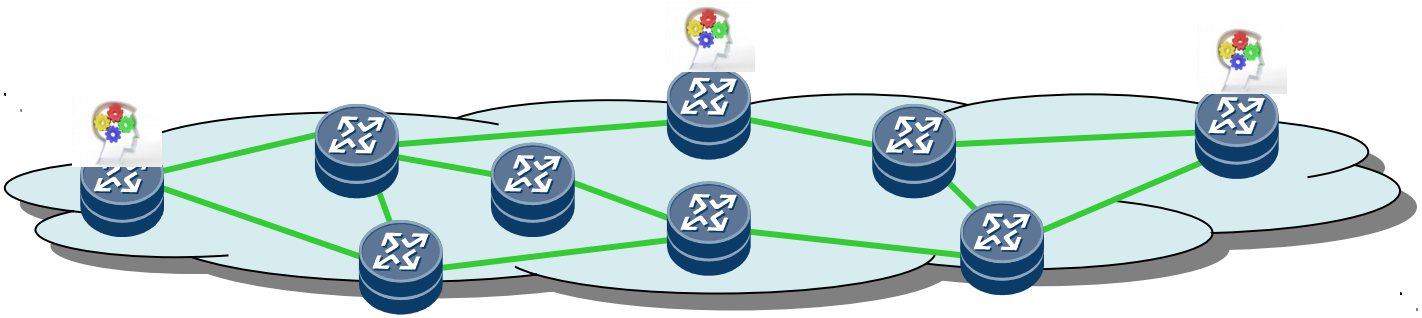
  #3 Single router learning

  #4 Distributed learning (Ensemble Learning)

- This is our focus
  - Router is the “first hand” entity to handle the traffic
    - Potential ***fast-loop*** to act on the traffic
  - Routers can sense the raw traffic continuously
    - Raw traffic might contain more potential valuable information co mparing to the extracted statistical/status data reported to the ce ntral node

#3: Single Router Learning

- One router only does its own learning task
  - Both training and prediction is done within the single router
- Only for router-local and lightweight prediction/decision
  - One router can only learn the single point knowledge
  - Restricted computing resource in routers



Use Case #3-1: Dynamic Traffic Alarm Threshold

- One router record traffic data in a per-interface m anner, attributes might contain:
  - Time stamp
  - Input traffic rate
  - Output traffic rate
- After a certain of time , the router learns the data:
  - Anomaly analysis of input/output traffic rate layout on the time period
  - To sort out the normal traffic range in a certain time pe riod

Use Case #3-1: Dynamic Traffic Alarm Threshold

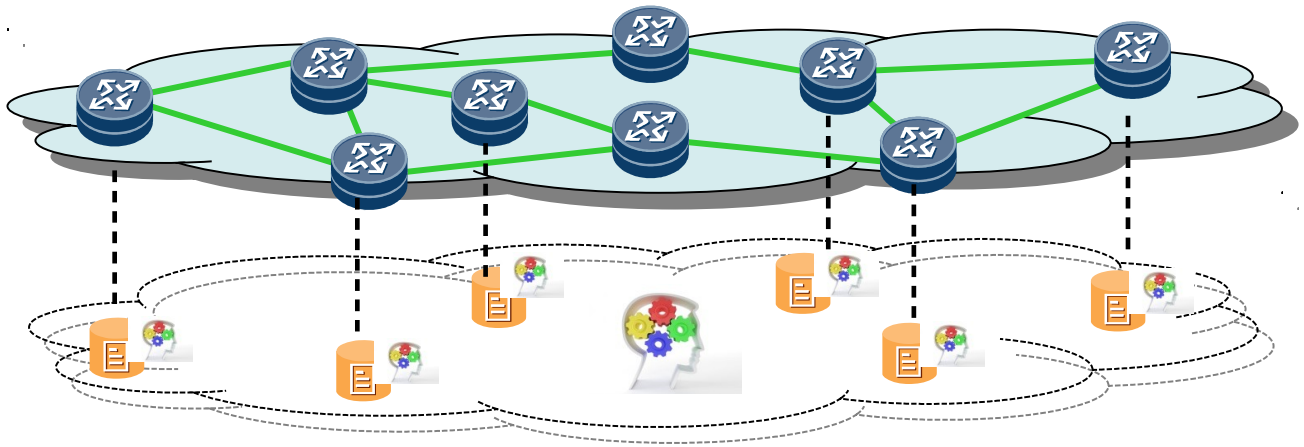
- Set up the threshold parameter for traffic alarm base d on the learnt results
- Dynamically adjust the threshold parameter based o n the new data
  - Periodical anomaly analysis
- When interface traffic exceeds the threshold
  - Report the event to the NMS
- Benefit
  - Faster detection than learning at the central repository
  - Abstracted report to the central repository, offload the ov erall burden of the network level learning

Use Case #3-2: Load Balancing

- Traditional scheduling algorithms
  - Random, Round Robin etc.
- ML-based scheduling
  - Making real-time test of some KPI as feedback
    - Server's load
    - Response time
    - Connections count

#4: Distributed Learning among Routers

- Distributed learning for the same task
  - Each router learns its own data/passing traffic ***(Local training)***
  - Ensemble all the routers' learning results ***(Global ensemble)***
  - Distribute the ensemble learning result to all routers ***(Global distribution)***
  - One router acts on the traffic independently ***(Local decision)***



Use Case #4-1: Distributed Traffic Anomaly D etection

- Separate the traffic analysis
  - For a core router or edge router, the traffic volume might be massive that it is difficult to do analysis on that router
  - The traffic is naturally separated to the downstream routers
- Anomaly detection on each downstream router
  - e.g. DDoS traffic detection
- Ensemble the detection result

Use Case #4-2: Dynamic Routing Policy

- Static path in a ring network
  - Always two alternative paths (CW&CCW) to achieve reliability
  - Traditional tech: one direction by default, the other for failover
- Dynamic path
  - Each router continuously sending probing packets to determine which direction is better
  - Ensemble each router's result, sort out the proper direction
  - All router in the ring switched to the new direction
  - This could be done periodically

(Pre) Perception

- In-router learning might leverage more on
  - Unsupervised learning
    - No labeled data for supervised learning in the traffic
  - Reinforcement learning
    - Improve the learning task performance based on autonomic real-time network testing/monitoring (Reinforcement learning)
    - More sophisticated autonomic testing technologies are needed
- Single router learning (#3) could be essential even there is central learning
- Distributed learning (#4) has the potential capability for autonomic optimization in a subset of routers (e.g., grouped by routing domain or path)

Future Plan

- Implementation/test data of the proposed use cases
- Explore Reinforcement Learning
- Router computational capability survey
  - What kind of learning task could be afforded by ro uters
  - Specifically, the resource and time consumption of the proposed use cases

Comments?

Thank you!

leo.liubing@huawei.com

IETF95, Buenos Arise